資料結構(Data Structures)

Course 3: Array (陣列)

授課教師:陳士杰

國立聯合大學 資訊管理學系

Outlines

- 本章重點
  - Array的定義
  - Array中元素儲存位置的計算
  - 多項式的表示
  - Sparse Matrix的表示
  - 特殊矩陣之儲存位置計算
    - 上、下三角矩陣
    - 對稱矩陣

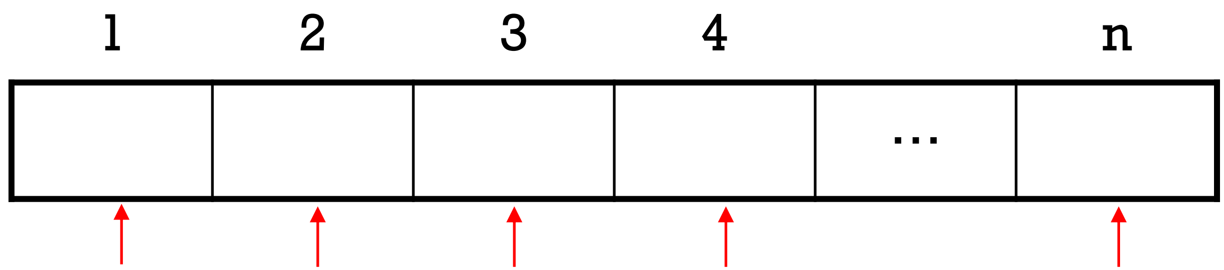
Array

- Def:
  - 為表示**有序串列**之一種方式
  - 其佔用**連續性的記憶體空間**
  - 各元素型態皆需**相同** (一致)
  - 支援**Sequential**及**Random Access**
  - **插入**、**刪除**元素較為麻煩
    - ∵需**挪移其它元素**
    - ∴不易動態增刪空間大小
    - Time = O(n)

說明

- 以插入一元素為例。

需挪動的格數:



所欲插入的位置:

$\therefore$ **平均挪移次數** $= (n+(n-1)+\ldots+1)/n = n(n-1)/2 \times 1/n$

$= (n+1)/2 =$ **$O(n)$**

Array中元素儲存位置的計算

- 一維陣列
- 二維陣列
- 三維陣列
- N維陣列

一維陣列

- **<u>宣告方式 1</u>:**
  - A[1 … n]表示有n格元素 (有些書表示為A[1:n])
  - 假設:
    - l_0表示Array的起始位址
    - d表示元素大小

⇨A[i]的location = l_0 + (i-1) × d

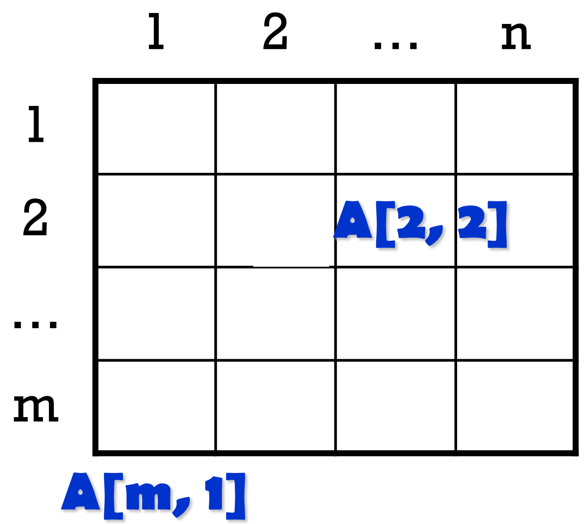
- **<u>宣告方式 2</u>:** **(一般化宣告, 起始位址不一定為 1)**
  - A[l … u]表示有 u-l+1 格元素(有些書表示為A[l:u])
  - 假設:
    - l_0表示Array的起始位址
    - d表示元素大小

⇨A[i]的location = l_0 + (i-l) × d

※範例練習※

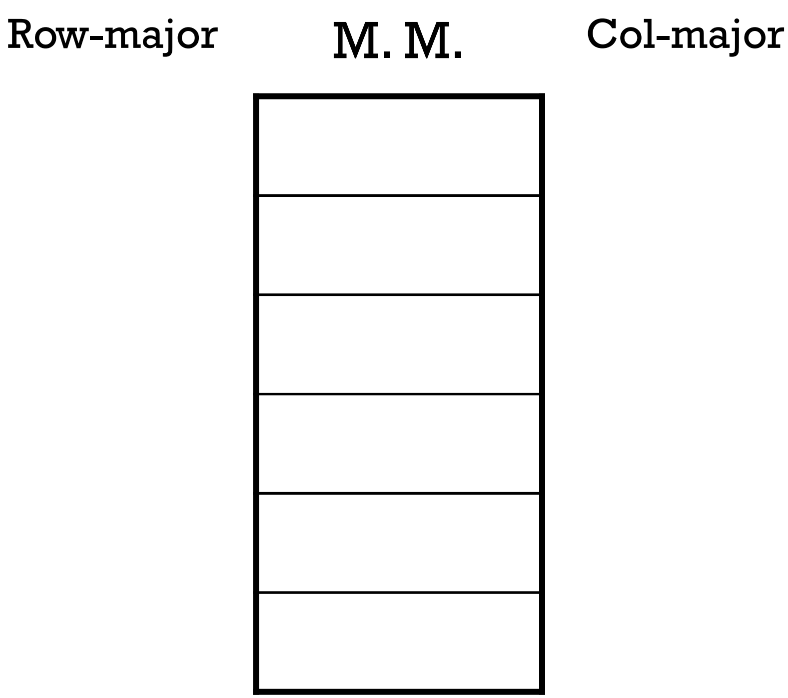
- 有一陣列A[0:100]，起始位置A[0] = 100，d=2，則A[16] = ?
  - Ans: 132
- 有一陣列A[-3:10]，起始位置A[-3] = 100，d=1，則A[5] = ?
  - Ans: 108

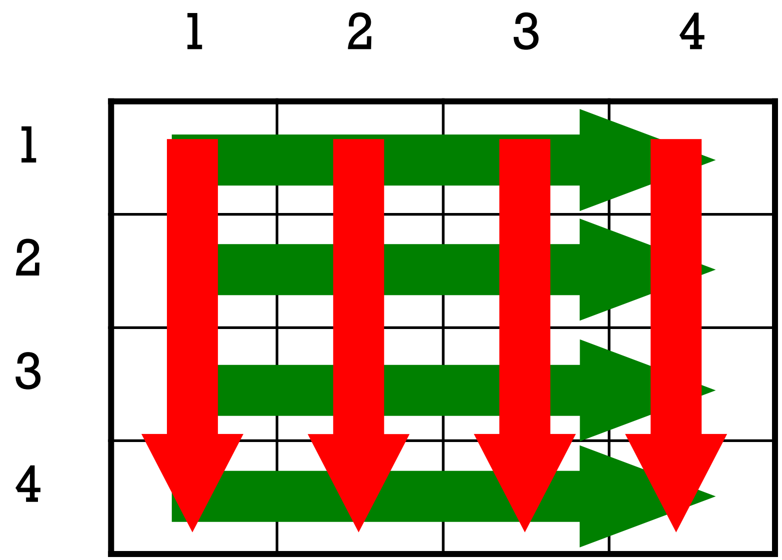
- **<u>宣告方式 1</u>**: A[1:m, 1:n] of data
  - 有 **<u>m</u>** 列 (Row), **<u>n</u>** 行 (Column), **<u>m×n</u>** 格
  - 假設:
    - l_0: 起始位址
    - d: 元素大小



二維陣列儲存在主記憶體中的方式有兩種:

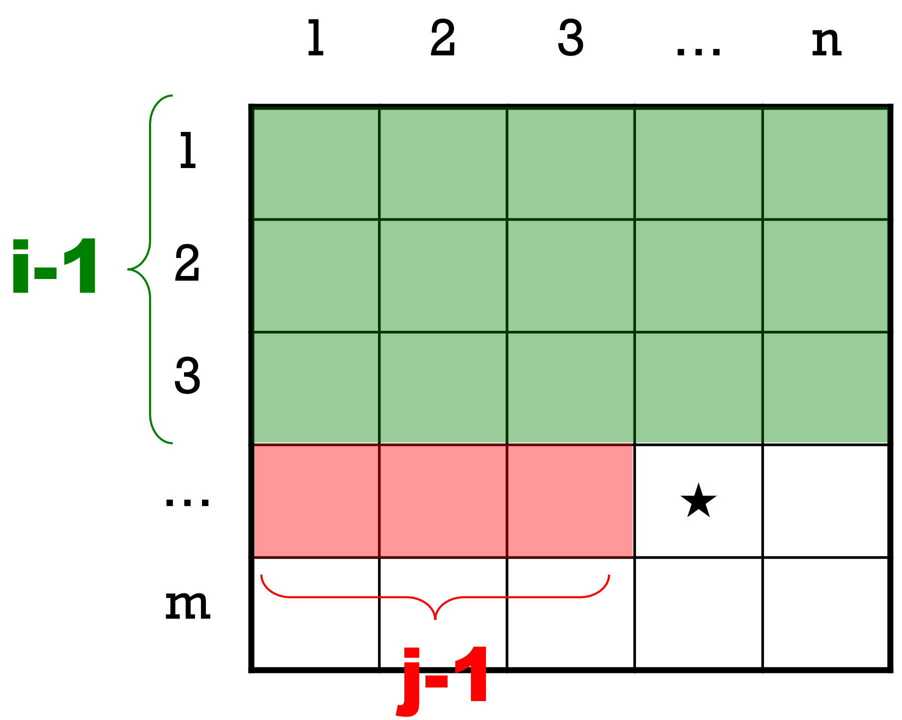
- Row-major
- Column-major





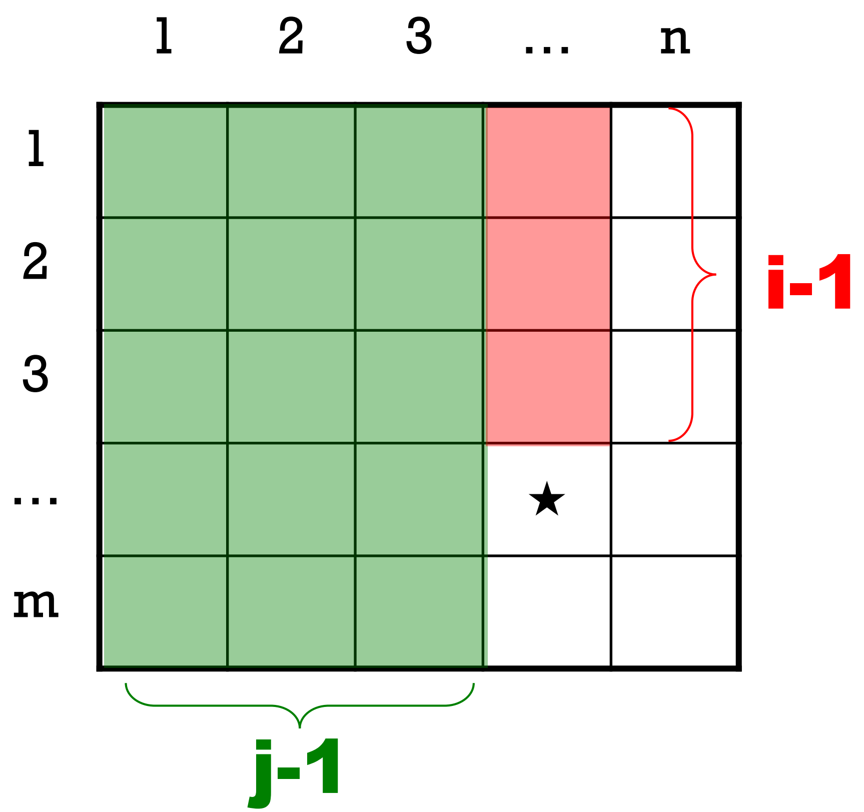
- Row-Major:

$$A[i, j] = l_0+[\qquad\qquad\qquad\qquad] \times d$$



- Column-Major:

$$A[i, j] = l_0 + [\qquad\qquad\qquad\qquad] \times d$$



- **宣告方式 2**: $A[l_1:u_1, l_2:u_2]$ of data
  - 有 $\mathbf{u_1-l_1+1}$ 列 (Row), $\mathbf{u_2-l_2+1}$ 行 (Column)
- Row-Major:

$$A[i, j] = l_0+[(i-l_1)\times(u_2-l_2+1) + (j-l_2)] \times d$$

- Column-Major:

$$A[i, j] = l_0+[(i-l_1) + (j-l_2) \times (u_1-l_1+1)] \times d$$

二維陣列所探討的4個議題

- 給予所有宣告，求A[i, j]之location
- 給予兩個已知位置，求A[i, j]之location
  - 須自行判斷出Row或Column-major
- 給予兩個已知位置，但判斷出Row或Column-major皆有可能，求A[i, j]之location
- 給予三個已知位置量，求A[i, j]之location

議題1:給予所有宣告, 求A[i, j]之location

- 例: A[-4:3, -3:2], 元素大小為 1, l_0 = 100, 採Row-major儲存, 求A[1, 1]之location.

Sol:

∵ 有 8 列, 有 6 行

∴ $A[1,1] = l_0 + [(1-(-4))\times 6 + (1-(-3))]\times 1$

$= 100 + [30+4] \times 1$

$=134$

※範例練習※

- 有一陣列A[-100:1, 1:100], 起始位置為100, d=4, 為Row-Major儲存, 則A[1, 12] = ?
  - Ans: 40544
- 有一陣列A[5:10, -10:20], 起始位置為100, d=4, 為Row-Major儲存, 則A[5, -5] = ?
  - Ans: 160

議題2：給予兩個已知位置，求A[i, j]之location

- 須自行判斷出**Row或Column-major**
- 例: 若A[3,2]之location為1110, A[2,3]之location為 1115, 元素大小為 1, 求 A[1, 4], A[5, 4]之location.

Sol:

1. 先判斷是Row-major或是Column-major

$$A[3, 2] = 1110$$

$$\vee \; \wedge \qquad \wedge$$

$$A[2, 3] = 1115$$

2. A[3,2]與A[2,3]代入公式，求解l_0與列數m

$$A[i, j] = l_0+[\mathbf{(i-1)} + \mathbf{(j-1)} \times \mathbf{m}] \times d$$

3. 求解出A[1, 4] = 1120, A[5, 4] = 1124

※範例練習※

- Assume in a byte machine. A is an array declared as A[-1:m, 2:n] and each element occupied 3 bytes. The address of A[3,5] is at 180 and A[5,3] is at 138. Find the address of the element A[-1,2].
  - Ans: 96

議題3:給予兩個已知位置, 但判斷出Row或Column-major皆有可能, 求A[i, j]之location.

- 算出的結果可能有以下兩種情況:
  - 1合, 1不合! (即: 可判斷出是哪一個major)
  - 兩者皆合! (兩個major的結果都要算)
- 例: 若A[3,3]之location為121, A[6,4]之location為159, 元素大小為 1, 求 A[4, 5]之location.

Sol:

$$A[3, 3] = 121$$

$$\wedge\ \wedge \quad \wedge$$

$$A[6, 4] = 159$$

1) 先算Row-Major:

- 將A[3,3]與A[6,4]代入以下公式，求解l_0與行數n:

$$A[i, j] = l_0+[(i-1)\times n + (j-1)] \times d$$

⇨ 得: n = 37/3 (∵非整數, ∴ 不是Row-Major)

2) 算Column-Major:

- 將A[3,3]與A[6,4]代入以下公式，求解l_0與列數m:

$$A[i, j] = l_0+[(i-1) + (j-1) \times m] \times d$$

⇨ 得: m = 35, l_0 = 49 (∵皆為整數, ∴ 是Column-Major)

⇨ 得: A[4, 5] = 192.

議題4:給予三個已知位置，求A[i, j]之location.

- 一定要假設元素大小 d (此時的元素大小不一定為 1)
- 當判斷出是Row-major或是Column-major時:
  - Row-major: 起始位置 l_0, n, d
  - Column-major: 起始位置 l_0, m, d
- 例:A(1, 1)之address為2，A(2, 3)為18，A(3, 2)為28，求A(4, 5)之Location?

Sol:

- 由 A(2, 3) 與 A(3, 2) 及其兩個位置，可以判斷出是Row-major (不要用A(1, 1)和另兩者之一比，會不好判斷!!)
- 此時，假設元素大小為 d，行數為 n

- $A(2,3) = A(1,1) + [(2-1) \times n + (3-1)] \times d$

  ⇨ $18 = 2 + nd + 2d$ ....❶

- $A(3,2) = A(1,1) + [(3-1) \times n + (2-1)] \times d$

  ⇨ $28 = 2 + 2nd + d$ .... ❷

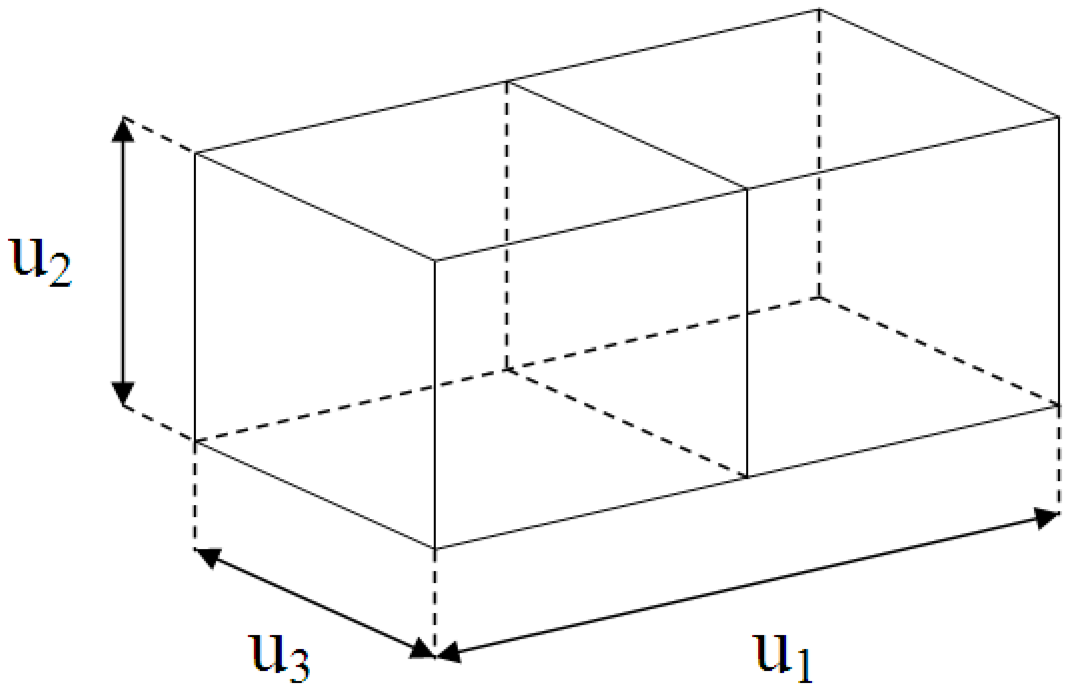
- 因此，❶ $\times 2$ - ❷ ⇨ $8 = 2 + 3d$，$\therefore$ $\underline{d = 2}$。

- 再將 d 代回公式 ❶，可得 $18 = 2 + 2n + 4$，$\therefore \underline{n = 6}$。

- $\therefore A(4,5) = A(1,1) + [(4-1) \times 6 + (5-1)] \times 2 =$ **46**。

三維陣列

- 從二維陣列起，不論是幾維陣列，在Computer中的陣列儲存於主記憶體的方式只有兩種:
  - Row-major
  - Column-major
- 三維陣列示意圖:



- 宣告: A[1:u_1, 1:u_2, 1:u_3], 起始位址為: l_0, 元素大小: d
  - 在Row-major下，A[i, j, k]之location為何?

$$\mathbf{A[i, j, k] = l_0 + [(i-1)\times u_2 u_3 + (j-1)\times u_3 + (k-1)] \times d}$$

(有u_1個二維陣列，每一個二維陣列皆是$u_2 \times u_3$這麼大)

  - 在Column-major下，A[i, j, k]之location為何?

$$\mathbf{A[i, j, k] = l_0 + [(i-1) + (j-1)\times u_1 + (k-1)\times u_2 u_1] \times d}$$

(有u_3個二維陣列，每一個二維陣列皆是$u_2 \times u_1$這麼大)

- 一般化宣告: $A[l_1:u_1, l_2:u_2, l_3:u_3]$
  - 在Row-major下，A[i, j, k]之location為何?

$$A[i, j, k] = l_0 + [(i-l_1)\times(u_2-l_2+1)\times(u_3-l_3+1) + (j-l_2)\times(u_3-l_3+1)+(k-l_3)] \times d$$

  - 在Column-major下，A[i, j, k]之location為何?

$$A[i, j, k] = l_0 + [(i-l_1)+(j-l_2)\times(u_1-l_1+1) +(k-l_3)\times(u_2-l_2+1)\times(u_1-l_1+1)] \times d$$

※範例練習※

- 有一陣列A[-3:2, -2:3, 0:4]，起始位置為 318，d=1，則分別求算在Row-major與Column-major下, A[1, 3, 3] 之 location？
  - Ans: (Row-major) 466; (Column-major) 460

多項式的表示方式

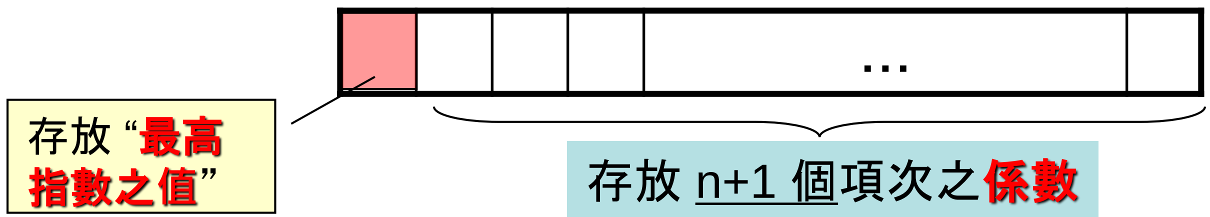
- 方法:
  - 按照指數由高到低，依序儲存係數
  - 儲存非零項次的係數與指數
  - 利用單向鏈結串列 (Single linked list)

用陣列表示

用鏈結串列表示

方法一

- 按照指數由高到低，依序儲存係數。此表示方法僅適用於**一元多次**之多項式 (即：多項式中僅一種變數)。
  - 作法：假設多項式中的最高指數為 n，表示最多會有n+1項次，則準備一個size為n+2的一維陣列。



  - 例：多項式 $f(x) = 5x^4+8x^3+2x^2+6$，則：

4	5	8	2	0	6

  - 缺點：當指數很大，且零項次極多時不適用。(∵太浪費空間)
  - 例：$3x^{100}+6 \Rightarrow$ 需準備 102 個儲存格，實際有用僅 3 格，浪費 99 格用於填 0

方法二

- 儲存非零項次的係數與指數。
  - 作法：假設在多項式中，有k個非零項次，則需準備一個size為<u>2k+1</u>的一維陣列。

1	2	3	4	5	6	7	…	2k+1
							…	

存放 "**非零項次之個數k**"

成對存放 k 個項次之**係數**與**指數**

  - 例：$3x^{100}+6 \Rightarrow$ 準備一個size為 $2\times2+1=5$ 的一維陣列來存放。

2	3	100	6	0

係數 指數 係數 指數

  - 缺點：不適用於非零項次極多的多項式，否則此法所使用的儲存空間大約是方法一的兩倍。

稀疏矩陣 (Sparse Matrix) 的表示方式

- Def: 非零元素個數極少的矩陣

$$
\begin{array}{c} \\ 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \end{array}
\begin{array}{c}
\begin{array}{cccccc} 1 & 2 & 3 & 4 & 5 & 6 \end{array} \\
\begin{bmatrix}
15 & 0 & 0 & 22 & 0 & 15 \\
0 & 11 & 3 & 0 & 0 & 0 \\
0 & 0 & 0 & 6 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 \\
91 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 28 & 0 & 0 & 0
\end{bmatrix}
\end{array}
$$

- 表示方式:
  - 利用一個m×n的二維陣列，一一對應矩陣之資料並儲存之。
    - 缺點: 極浪費空間 (如上例，存了一堆零!!)
  - 利用3-tuple來儲存非零元素。

利用3-tuple來儲存非零元素

- 3-tuple結構：以 <列, 行, 值> 來表示一個非零元素。
  - 假設 m×n 的稀疏矩陣，有 k 個非零元素，則準備一個size為(k+1)×3的二維陣列。

存放 3-tuple 矩陣的列數 m，行數 n，及非零的個數k

	1	2	3
0	m	n	K
1			
2	<列	行	值>
…		…	
K			

存放 k 個非零值之 Tuple 值 (rows 1 … K)

  - 以Row-major方式依序存放資料

請將前述之稀疏矩陣以3-tuple存放。

Ans:

- 前述之稀疏矩陣大小為 6×6，其中有8個非零元素。故需準備一個size為(8+1)×3 的二維陣列如下：

	1	2	3
0	**6**	**6**	**8**
1	1	1	15
2	1	4	22
3	1	6	15
4	2	2	11
5	2	3	3
6	3	4	6
7	5	1	91
8	6	3	28

特殊矩陣的元素儲存位置計算

- 上、下三角矩陣 (Upper, Lower Triangular Matrix)
- 對稱矩陣 (Symmetric Matrix)

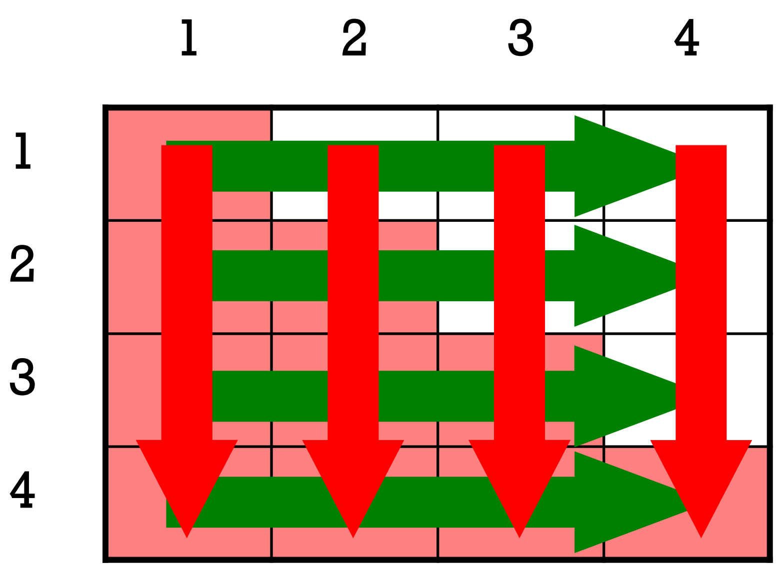
下三角矩陣 (Lower Triangular Matrix)

- Def:
  - 為一個n × n矩陣，其對角線以上的元素均為零
  - 為一個n × n矩陣，其中元素 $a_{ij} = 0$, if $i<j$

$$\begin{array}{cc} & \begin{array}{cccccc} 1 & 2 & 3 & \cdot\cdot & \cdot\cdot & n \end{array} \\ \begin{array}{c} 1 \\ 2 \\ 3 \\ : \\ : \\ n \end{array} & \begin{bmatrix} \times & & & & & \\ \times & \times & & & 0 & \\ \times & \times & \times & & & \\ \times & \times & \times & \times & & \\ \times & \times & \times & \times & \times & \\ \times & \times & \times & \times & \times & \times \end{bmatrix} \end{array}$$

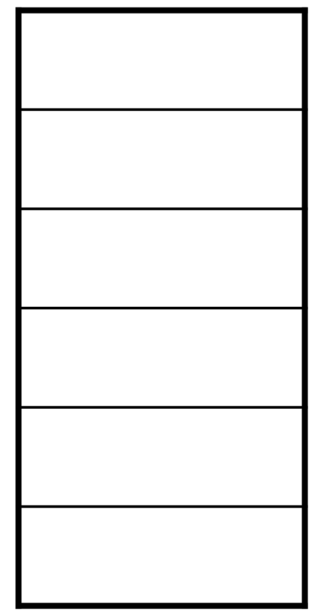
- (非零) 元素個數 : $\underline{n(n+1)/2}$ $\Rightarrow \because 1+2+3+\ldots+n = n(n+1)/2$ (以**列**來算)
- 零元素個數 : $\underline{n^2 - n(n+1)/2}$

- 欲將下三角矩陣存放到一維陣列 B:
  - 此一維陣列的Size應為: n(n+1)/2
  - 存放方式:
    - Row-major
    - Column-major



Row-major

B[k]



Col-major

- Row-major

k = ______________ + ____________ = ____________

前(i-1)列共有？格元素　　第 i 列上的第?格

$$
\begin{array}{ccc}
 & & \begin{array}{cccccc} & & & j & & \\ 1 & 2 & 3 & \cdot\cdot & \cdot\cdot & n \end{array} \\
 & \begin{array}{c} 1 \\ 2 \\ 3 \\ : \\ : \\ n \end{array} & \left[\begin{array}{cccccc} \times & & & & & \\ \times & \times & & & & \\ \times & \times & \times & & & \\ \times & \times & \times & \times & & \\ \times & \times & \times & \times & \times & \\ \times & \times & \times & \times & \times & \times \end{array}\right] \\
\end{array}
$$

i →

- Column-major

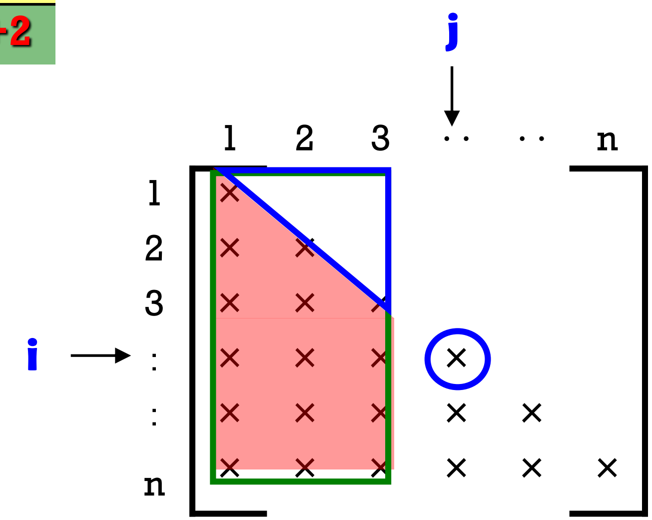
k = ______________ + ______________

前(j-1)行共有？格元素　　第 j 行上的第?格

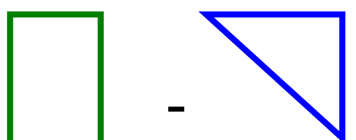
- 公式求法 1:

行數	1	2	3	…	j-1
元素個數	n	n-1	n-2	…	n-(j-1)+1 = **n-j+2**

$$
\begin{aligned}
k &= [(n-j+2)+n](j-1)/2 + (i-j+1) \\
&= [(2n-j+2-2)(j-1)]/2 + i \\
&= n(j-1) - j(j-1)/2 + i
\end{aligned}
$$



- 公式求法 2:



(空格數)

$$
\begin{aligned}
k &= \underline{\qquad \text{梯 形} \qquad} + \underline{\qquad i-j+1 \qquad} \\
&= [\mathbf{n(j-1)} - \mathbf{(0+1+2+3+\ldots+(j-2))}] + [i-(j-1)] \\
&= n(j-1) - (0+1+2+3+\ldots+(j-2)+(j-1)) + i \\
&= n(j-1) - j(j-1)/2 + i
\end{aligned}
$$

上三角矩陣 (Upper Triangular Matrix)

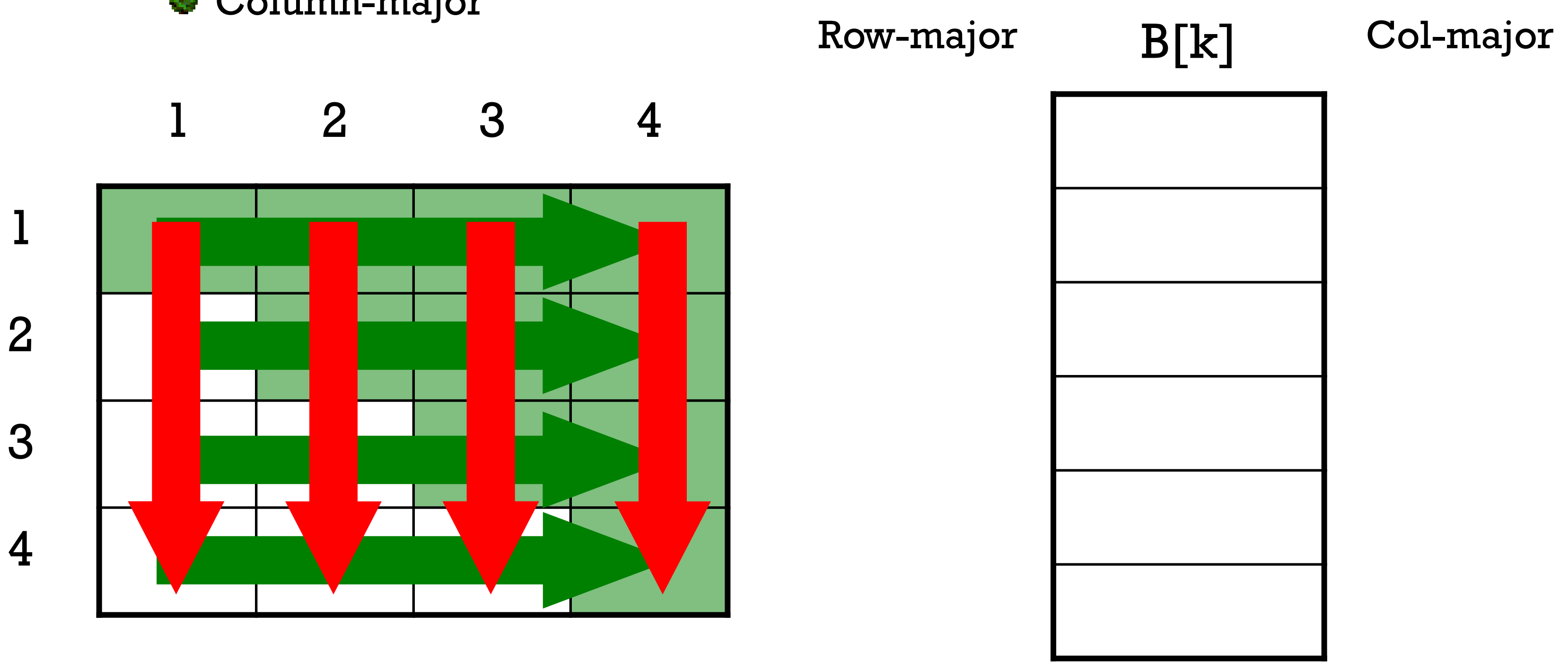
- Def:
  - 為一個n×n矩陣，其對角線以下的元素均為零
  - 為一個n×n矩陣，其中元素 $a_{ij} = 0$, if i>j

$$
\begin{array}{cc}
 & \begin{array}{cccccc} 1 & 2 & 3 & \cdot\cdot & \cdot\cdot & n \end{array} \\
\begin{array}{c} 1 \\ 2 \\ 3 \\ : \\ : \\ n \end{array} &
\begin{bmatrix}
\times & \times & \times & \times & \times & \times \\
 & \times & \times & \times & \times & \times \\
 & & \times & \times & \times & \times \\
 & & & \times & \times & \times \\
 & 0 & & & \times & \times \\
 & & & & & \times
\end{bmatrix}
\end{array}
$$

- (非零) 元素個數：<u>n(n+1)/2</u>

$\Rightarrow \because$ 1+2+3+…+n = n(n+1)/2 (以**行**來算)

- 欲將上三角矩陣存放到一維陣列 B:
  - 此一維陣列的Size應為: n(n+1)/2
  - 存放方式:
    - Row-major
    - Column-major



上三角矩陣A的值a_{ij}欲存入一維陣列B[k]，k=?

- Column-major
  - 分析方法與下三角矩陣之<u>Row-major</u>公式相同，故其分析結果與下三角矩陣的Row-major公式類似，只是<u>公式中的 i 與 j 互換</u>。

前(j-1)行共有？格元素　　第 j 行上的第?格

$$k = \underline{(1+2+3+\ldots+(j-1))} + \underline{\quad i \quad}$$

$$= \underline{j \times (j-1)/2 + i}$$

$$\begin{array}{c|cccccc|}
 & 1 & 2 & 3 & \cdot\cdot\,(j) & \cdot\cdot & n \\
\hline
1 & \times & \times & \times & \times & \times & \times \\
2 & & \times & \times & \times & \times & \times \\
3 & & & \times & \times & \times & \times \\
i \rightarrow \ : & & & & \times & \times & \times \\
: & & & & & \times & \times \\
n & & & & & & \times \\
\hline
\end{array}$$

- Row-major
  - 分析方法與下三角矩陣之<u>Column-major</u>公式相同，故其分析結果與下三角矩陣的Column-major公式類似，只是<u>公式中的 i 與 j 互換</u>。

前(i-1)列共有？格元素　　第 i 列上的第?格

$$k = \underline{\quad 梯形 \quad} + \underline{\quad j-i+1 \quad}$$

$$= n(i-1) - i(i-1)/2 + j$$

※範例練習※

- 例：有一個$A_{100\times100}$的下三角矩陣，回答下列問題：
  - 元素個數?
  - A[65, 60]的資料以Row-major存入一維陣列B的第k格 (B的起始格為1)，則k = ?
  - A[70, 50]以Column-major存入B[k]中，則k=?

Ans:<u>5050</u>, <u>2140</u>, <u>3745</u>

- Def: 對稱矩陣為一個 $n \times n$ 的矩陣，其 $a_{ij} = a_{ji}$
  - 例：

$$
\begin{array}{c c}
 & \begin{array}{ccc} 1 & 2 & 3 \end{array} \\
\begin{array}{c} 1 \\ 2 \\ 3 \end{array} & \begin{bmatrix} 4 & 1 & -5 \\ 1 & 7 & 2 \\ -5 & 2 & 8 \end{bmatrix}
\end{array}
$$

  - 較節省的儲存方式：只存上 (或下) 三角部份即可
    - 所需空間：$n(n+1)/2$

- A symmetric matrix is stored with its upper-diagonal part column-wisely on a one dimensional array, i.e., a_{11} stored in A(1); $a_{12}=a_{21}$ stored in A(2); a_{22} in A(3); $a_{31}=a_{13}$ in A(4); $a_{23}=a_{32}$ in A(5) and so on. Let a_{ij} be stored in A(k). Write a single expression for k in terms of i and j. The MAX and MIN function can be used in the expression.

Ans:

- 題意：

	1	2	3
1	4	1	-5
2	1	7	2
3	-5	2	8

⇨

	1	2	3	4	5	6
A	4	1	7	-5	2	8

- 本題在儲存資料時僅使用Column-major。但是，我們是否可以設計出一個單一的表示式，可任意於對稱矩陣中指定一組(i, j)資料，將所指定的資料置於一維陣列 A 當中、合乎原題目要求之所在位置 k。

如何將資料置於合乎原題意之一維陣列 A 之位置 k？

- 若資料 a_{ij} 位於**上三角** (即：i≤j)，則以Column-major的方式計算 k。

$$k = \frac{j \times (j-1)}{2} + i, \quad \text{if } i \le j.$$

	1	2	3
1	4	1	-5
2	1	7	2
3	-5	2	8

	1	2	3	4	5	6
A	4	1	7	-5	2	8

- 若資料a_{ij} 位於**下三角** (即：i≥j)，則以Row-major的方式計算 k。

$$k = \frac{i \times (i-1)}{2} + j, \quad \text{if } i \ge j.$$

	1	2	3
1	4	1	-5
2	1	7	2
3	-5	2	8

	1	2	3	4	5	6
A	4	1	7	-5	2	8

- 將前兩式合併成一個單一表示式:
  - 根據兩式的 if 判斷式，可以發現不論是在上三角公式還是下三角公式，在除法的分子部份使用的都是 i 與 j 當中最大那一項；而加法後面的項目則是 i 與 j 當中最小那一項。因此，藉由MAX與MIN的運算我們可以將兩式合併成一個單一表示式。

$$k=\frac{j\times(j-1)}{2}+i,\quad \text{if } i\le j.$$

$$+$$

$$k=\frac{i\times(i-1)}{2}+j,\quad \text{if } i\ge j.$$

$$\Rightarrow\quad k=\frac{MAX(i,\ j)\times(MAX(i,\ j)-1)}{2}+MIN(i,\ j).$$

  - 上述單一表示式可讓使用者從對稱矩陣中任挑一組資料(i, j)，放入到一維陣列A中，且其置放位址 k 可滿足原題意的要求。

補　充

N維陣列

- 宣告: A[1:u_1, 1:u_2, 1:u_3 , …, 1:u_n], 起始位址為: l_0, 元素大小: d
  - 在Row-major下:

$$A[i_1, i_2, \ldots, i_n] = l_0 + \begin{bmatrix} (i_1 - 1) \times u_2 u_3 \ldots u_n + \\ (i_2 - 1) \times u_3 u_4 \ldots u_n + \\ (i_n - 1) \times 1 \end{bmatrix} \times d$$

$$= l_0 + \left(\sum_{j=1}^{n} \left[(i_j - 1) \times \prod_{k=j+1}^{n} u_k \right] \right) \times d, \quad \text{且} \prod_{k=n+1}^{n} u_k = 1$$

在Column-major下:

$$A[i_1,i_2,\ldots,i_n]=l_0+\begin{bmatrix}(i_1-1)\times u_{n-1}u_{n-2}\ldots u_1+\\(i_2-1)\times u_{n-2}u_{n-3}\ldots u_1+\\(i_n-1)\times 1\end{bmatrix}\times d$$

$$=l_0+\left(\sum_{j=1}^{n}\left[(i_j-1)\times\prod_{k=1}^{j-1}u_k\right]\right)\times d,\quad \text{且}\prod_{k=1}^{0}u_k=1$$

範例練習

- 有一陣列A[-2:7, -4:10, -2:1 , -3:2 , 1:10], 起始位置為 38, d=8, 則求算在Row-major下, A[0, 8, 0, 1, 8] 之location？
  - Ans: 82014